

PAGE: 1

RAW SEQUENCE LISTING
PATENT APPLICATION US/09/175,683ADATE: 01/05/2000
TIME: 15:11:36

Input Set: I175683A.RAW

<p>This Raw Listing contains the General Information Section and up to first 5 pages.</p>

ENTERED

```

1  <110> APPLICANT: Chen, Li How
2      Meade, Henry
3  <120> TITLE OF INVENTION: NOVEL MODIFIED MSP-1 NUCLEIC ACID SEQUENCES AND METHODS
4      FOR INCREASING mRNA LEVELS AND PROTEIN EXPRESSIONS IN
5      CELL SYSTEMS
6  <130> FILE REFERENCE: 10275/134001
7  <140> CURRENT APPLICATION NUMBER: US/09/175,683A
8  <141> CURRENT FILING DATE: 1998-10-20
9  <150> EARLIER APPLICATION NUMBER: US 60/095,649
10 <151> EARLIER FILING DATE: 1998-08-07
11 <150> EARLIER APPLICATION NUMBER: US 60/062,592
12 <151> EARLIER FILING DATE: 1997-10-20
13 <160> NUMBER OF SEQ ID NOS: 13
14 <170> SOFTWARE: FastSEQ for Windows Version 4.0
15 <210> SEQ ID NO 1
16 <211> LENGTH: 1065
17 <212> TYPE: DNA
18 <213> ORGANISM: preferably, a bacterium, virus, or parasite
19 <220> FEATURE:
20 <221> NAME/KEY: CDS
21 <222> LOCATION: (1)...(1065)
22 <400> SEQUENCE: 1
23      gcc gtc act ccc tcc gtc atc gat aac atc ctg tcc aag atc gag aac      48
24      Ala Val Thr Pro Ser Val Ile Asp Asn Ile Leu Ser Lys Ile Glu Asn
25      1          5          10          15
26      gag tac gag gtg ctg tac ctg aag ccg ctg gca ggg gtc tac cgg agc      96
27      Glu Tyr Glu Val Leu Tyr Leu Lys Pro Leu Ala Gly Val Tyr Arg Ser
28      20          25          30
29      ctg aag aag cag ctg gag aac aac gtg atg acc ttc aac gtg aac gtg      144
30      Leu Lys Lys Gln Leu Glu Asn Asn Val Met Thr Phe Asn Val Asn Val
31      35          40          45
32      aag gat atc ctg aac agc cgg ttc aac aag cgg gag aac ttc aag aac      192
33      Lys Asp Ile Leu Asn Ser Arg Phe Asn Lys Arg Glu Asn Phe Lys Asn
34      50          55          60
35      gtg ctg gag agc gat ctg atc ccc tac aag gat ctg acc agc agc aac      240
36      Val Leu Glu Ser Asp Leu Ile Pro Tyr Lys Asp Leu Thr Ser Ser Asn
37      65          70          75          80
38      tac gtg gtc aag gat ccc tac aag ttc ctg aac aag gag aag aga gat      288
39      Tyr Val Val Lys Asp Pro Tyr Lys Phe Leu Asn Lys Glu Lys Arg Asp
40      85          90          95
41      aag ttc ctg agc agt tac aac tac atc aag gat agc att gat acc gat      336
42      Lys Phe Leu Ser Ser Tyr Asn Tyr Ile Lys Asp Ser Ile Asp Thr Asp
43      100         105         110
44      atc aac ttc gcc aac gat gtc ctg gga tac tac aag atc ctg tcc gag      384

```

PAGE: 2

RAW SEQUENCE LISTING
PATENT APPLICATION US/09/175,683A

DATE: 01/05/2000
TIME: 15:11:36

Input Set: I175683A.RAW

45	Ile Asn Phe Ala Asn Asp Val Leu Gly Tyr Tyr Lys Ile Leu Ser Glu	
46	115 120 125	
47	aag tac aag agc gat ctg gat tca atc aag aag tac atc aac gat aag	432
48	Lys Tyr Lys Ser Asp Leu Asp Ser Ile Lys Lys Tyr Ile Asn Asp Lys	
49	130 135 140	
50	cag gga gag aac gag aag tac ctg ccc ttc ctg aac aac atc gag acc	480
51	Gln Gly Glu Asn Glu Lys Tyr Leu Pro Phe Leu Asn Asn Ile Glu Thr	
52	145 150 155 160	
53	ctg tac aag acc gtc aac gat aag att gat ctg ttc gtg atc cac ctg	528
54	Leu Tyr Lys Thr Val Asn Asp Lys Ile Asp Leu Phe Val Ile His Leu	
55	165 170 175	
56	gag gcc aag gtc ctg aac tac aca tat gag aag agc aac gtg gag gtc	576
57	Glu Ala Lys Val Leu Asn Tyr Thr Tyr Glu Lys Ser Asn Val Glu Val	
58	180 185 190	
59	aag atc aag gag ctg aat tac ctg aag acc atc cag gat aag ctg gcc	624
60	Lys Ile Lys Glu Leu Asn Tyr Leu Lys Thr Ile Gln Asp Lys Leu Ala	
61	195 200 205	
62	gat ttc aag aag aac aac aac ttc gtc ggg atc gcc gat ctg agc acc	672
63	Asp Phe Lys Lys Asn Asn Asn Phe Val Gly Ile Ala Asp Leu Ser Thr	
64	210 215 220	
65	gat tac aac cac aac aac ctg ctg acc aag ttc ctg agc acc ggt atg	720
66	Asp Tyr Asn His Asn Asn Leu Leu Thr Lys Phe Leu Ser Thr Gly Met	
67	225 230 235 240	
68	gtc ttc gaa aac ctg gcc aag acc gtc ctg agc aac ctg ctg gat ggg	768
69	Val Phe Glu Asn Leu Ala Lys Thr Val Leu Ser Asn Leu Leu Asp Gly	
70	245 250 255	
71	aac ctg cag ggg atg ctg aac atc agc cag cac cag tgt gtg aag aag	816
72	Asn Leu Gln Gly Met Leu Asn Ile Ser Gln His Gln Cys Val Lys Lys	
73	260 265 270	
74	cag tgt ccc cag aac agc ggg tgt ttc aga cac ctg gat gag aga gag	864
75	Gln Cys Pro Gln Asn Ser Gly Cys Phe Arg His Leu Asp Glu Arg Glu	
76	275 280 285	
77	gag tgt aag tgt ctg ctg aac tac aag cag gaa ggt gat aag tgt gtg	912
78	Glu Cys Lys Cys Leu Leu Asn Tyr Lys Gln Glu Gly Asp Lys Cys Val	
79	290 295 300	
80	gaa aac ccc aat cct act tgt aac gag aac aat ggt gga tgt gat gcc	960
81	Glu Asn Pro Asn Pro Thr Cys Asn Glu Asn Asn Gly Gly Cys Asp Ala	
82	305 310 315 320	
83	gat gcc aag tgt acc gag gag gat tca ggg agc aac ggg aag aag atc	1008
84	Asp Ala Lys Cys Thr Glu Glu Asp Ser Gly Ser Asn Gly Lys Lys Ile	
85	325 330 335	
86	acc tgt gag tgt acc aag cct gat tct tat cca ctg ttc gat ggt atc	1056
87	Thr Cys Glu Cys Thr Lys Pro Asp Ser Tyr Pro Leu Phe Asp Gly Ile	
88	340 345 350	
89	ttc tgt agt	1065
90	Phe Cys Ser	
91	355	
92	<210> SEQ ID NO 2	
93	<211> LENGTH: 1088	
94	<212> TYPE: DNA	

PAGE: 3

RAW SEQUENCE LISTING
PATENT APPLICATION US/09/175,683A

DATE: 01/05/2000
TIME: 15:11:36

Input Set: I175683A.RAW

```

95 <213> ORGANISM: preferably, a bacterium, virus, or parasite
96 <220> FEATURE:
97 <221> NAME/KEY: CDS
98 <222> LOCATION: (1)...(1083)
99 <400> SEQUENCE: 2
100   gca gta act cct tcc gta att gat aac ata ctt tct aaa att gaa aat      48
101   Ala Val Thr Pro Ser Val Ile Asp Asn Ile Leu Ser Lys Ile Glu Asn
102       1             5             10             15
103   gaa tat gag gtt tta tat tta aaa cct tta gca ggt gtt tat aga agt      96
104   Glu Tyr Glu Val Leu Tyr Leu Lys Pro Leu Ala Gly Val Tyr Arg Ser
105               20             25             30
106   tta aaa aaa caa tta gaa aat aac gtt atg aca ttt aat gtt aat gtt      144
107   Leu Lys Lys Gln Leu Glu Asn Asn Val Met Thr Phe Asn Val Asn Val
108               35             40             45
109   aag gat att tta aat tca cga ttt aat aaa cgt gaa aat ttc aaa aat      192
110   Lys Asp Ile Leu Asn Ser Arg Phe Asn Lys Arg Glu Asn Phe Lys Asn
111       50             55             60
112   gtt tta gaa tca gat tta att cca tat aaa gat tta aca tca agt aat      240
113   Val Leu Glu Ser Asp Leu Ile Pro Tyr Lys Asp Leu Thr Ser Ser Asn
114       65             70             75             80
115   tat gtt gtc aaa gat cca tat aaa ttt ctt aat aaa gaa aaa aga gat      288
116   Tyr Val Val Lys Asp Pro Tyr Lys Phe Leu Asn Lys Glu Lys Arg Asp
117               85             90             95
118   aaa ttc tta agc agt tat aat tat att aag gat tca ata gat acg gat      336
119   Lys Phe Leu Ser Ser Tyr Asn Tyr Ile Lys Asp Ser Ile Asp Thr Asp
120               100            105            110
121   ata aat ttt gca aat gat gtt ctt gga tat tat aaa ata tta tcc gaa      384
122   Ile Asn Phe Ala Asn Asp Val Leu Gly Tyr Tyr Lys Ile Leu Ser Glu
123       115            120            125
124   aaa tat aaa tca gat tta gat tca att aaa aaa tat atc aac gac aaa      432
125   Lys Tyr Lys Ser Asp Leu Asp Ser Ile Lys Lys Tyr Ile Asn Asp Lys
126       130            135            140
127   caa ggt gaa aat gag aaa tac ctt ccc ttt tta aac aat att gag acc      480
128   Gln Gly Glu Asn Glu Lys Tyr Leu Pro Phe Leu Asn Asn Ile Glu Thr
129       145            150            155            160
130   tta tat aaa aca gtt aat gat aaa att gat tta ttt gta att cat tta      528
131   Leu Tyr Lys Thr Val Asn Asp Lys Ile Asp Leu Phe Val Ile His Leu
132               165            170            175
133   gaa gca aaa gtt cta aat tat aca tat gag aaa tca aac gta gaa gtt      576
134   Glu Ala Lys Val Leu Asn Tyr Thr Tyr Glu Lys Ser Asn Val Glu Val
135       180            185            190
136   aaa ata aaa gaa ctt aat tac tta aaa aca att caa gac aaa ttg gca      624
137   Lys Ile Lys Glu Leu Asn Tyr Leu Lys Thr Ile Gln Asp Lys Leu Ala
138       195            200            205
139   gat ttt aaa aaa aat aac aat ttc gtt gga att gct gat tta tca aca      672
140   Asp Phe Lys Lys Asn Asn Asn Phe Val Gly Ile Ala Asp Leu Ser Thr
141       210            215            220
142   gat tat aac cat aat aac tta ttg aca aag ttc ctt agt aca ggt atg      720
143   Asp Tyr Asn His Asn Asn Leu Leu Thr Lys Phe Leu Ser Thr Gly Met
144       225            230            235            240

```

PAGE: 4

RAW SEQUENCE LISTING
PATENT APPLICATION US/09/175,683A

DATE: 01/05/2000
TIME: 15:11:36

Input Set: I175683A.RAW

```

145      gtt ttt gaa aat ctt gct aaa acc gtt tta tct aat tta ctt gat gga      768
146      Val Phe Glu Asn Leu Ala Lys Thr Val Leu Ser Asn Leu Leu Asp Gly
147              245              250              255
148      aac ttg caa ggt atg tta aac att tca caa cac caa tgc gta aaa aaa      816
149      Asn Leu Gln Gly Met Leu Asn Ile Ser Gln His Gln Cys Val Lys Lys
150              260              265              270
151      caa tgt cca caa aat tct gga tgt ttc aga cat tta gat gaa aga gaa      864
152      Gln Cys Pro Gln Asn Ser Gly Cys Phe Arg His Leu Asp Glu Arg Glu
153              275              280              285
154      gaa tgt aaa tgt tta tta aat tac aaa caa gaa ggt gat aaa tgt gtt      912
155      Glu Cys Lys Cys Leu Leu Asn Tyr Lys Gln Glu Gly Asp Lys Cys Val
156              290              295              300
157      gaa aat cca aat cct act tgt aac gaa aat aat ggt gga tgt gat gca      960
158      Glu Asn Pro Asn Pro Thr Cys Asn Glu Asn Asn Gly Gly Cys Asp Ala
159      305              310              315              320
160      gat gcc aaa tgt acc gaa gaa gat tca ggt agc aac gga aag aaa atc      1008
161      Asp Ala Lys Cys Thr Glu Glu Asp Ser Gly Ser Asn Gly Lys Lys Ile
162              325              330              335
163      aca tgt gaa tgt act aaa cct gat tct tat cca ctt ttc gat ggt att      1056
164      Thr Cys Glu Cys Thr Lys Pro Asp Ser Tyr Pro Leu Phe Asp Gly Ile
165              340              345              350
166      ttc tgc agt cac cac cac cac cac cac taact      1088
167      Phe Cys Ser His His His His His His
168              355              360
169      <210> SEQ ID NO 3
170      <211> LENGTH: 88
171      <212> TYPE: DNA
172      <213> ORGANISM: preferably, a bacterium, virus, or parasite
173      <400> SEQUENCE: 3
174      tcgacgagag ccatgaaggt cctcatcctt gcctgtcttg tggctctggc cattgcaaga      60
175      gagcaggaag aactcaatgt agtcggtg      88
176      <210> SEQ ID NO 4
177      <211> LENGTH: 88
178      <212> TYPE: DNA
179      <213> ORGANISM: preferably, a bacterium, virus, or parasite
180      <400> SEQUENCE: 4
181      gatctaccga ctacattgag ttcttctctgc tctcttgcaa tggccagagc caccagacag      60
182      gcaaggatga ggaccttcat ggctctcg      88
183      <210> SEQ ID NO 5
184      <211> LENGTH: 60
185      <212> TYPE: DNA
186      <213> ORGANISM: preferably, a bacterium, virus, or parasite
187      <400> SEQUENCE: 5
188      taactcgagc gaaccatgaa ggtcctcatc cttgcctgtc tgggtggctct ggccattgca      60
189      <210> SEQ ID NO 6
190      <211> LENGTH: 48
191      <212> TYPE: DNA
192      <213> ORGANISM: preferably, a bacterium, virus, or parasite
193      <400> SEQUENCE: 6
194      aattctcgag ttagtggtgg tgggtgggtg gactgcagaa ataccatc      48

```

PAGE: 5

RAW SEQUENCE LISTING
PATENT APPLICATION US/09/175,683A

DATE: 01/05/2000
TIME: 15:11:36

Input Set: I175683A.RAW

```

195 <210> SEQ ID NO 7
196 <211> LENGTH: 31
197 <212> TYPE: DNA
198 <213> ORGANISM: preferably, a bacterium, virus, or parasite
199 <400> SEQUENCE: 7
200      aatagatctg cagtaactcc ttccgtaatt g                               31
201 <210> SEQ ID NO 8
202 <211> LENGTH: 1142
203 <212> TYPE: DNA
204 <213> ORGANISM: preferably, a bacterium, virus, or parasite
205 <220> FEATURE:
206 <221> NAME/KEY: CDS
207 <222> LOCATION: (1)...(1128)
208 <400> SEQUENCE: 8
209      atg aag gtc ctc ata att gcc tgt ctg gtg gct ctg gcc att gca gcc           48
210      Met Lys Val Leu Ile Ile Ala Cys Leu Val Ala Leu Ala Ile Ala Ala
211           1             5             10             15
212      gtc act ccc tcc gtc atc gat aac atc ctg tcc aag atc gag aac gag           96
213      Val Thr Pro Ser Val Ile Asp Asn Ile Leu Ser Lys Ile Glu Asn Glu
214           20             25             30
215      tac gag gtg ctg tac ctg aag ccc ctg gca gga gtc tac agg agc ctg          144
216      Tyr Glu Val Leu Tyr Leu Lys Pro Leu Ala Gly Val Tyr Arg Ser Leu
217           35             40             45
218      aag aag cag ctg gag aac aac gtg atg acc ttc aac gtg aac gtg aag          192
219      Lys Lys Gln Leu Glu Asn Asn Val Met Thr Phe Asn Val Asn Val Lys
220           50             55             60
221      gat atc ctg aac agc agg ttc aac aag agg gag aac ttc aag aac gtg          240
222      Asp Ile Leu Asn Ser Arg Phe Asn Lys Arg Glu Asn Phe Lys Asn Val
223           65             70             75             80
224      ctg gag agc gat ctg atc ccc tac aag gat ctg acc agc agc aac tac          288
225      Leu Glu Ser Asp Leu Ile Pro Tyr Lys Asp Leu Thr Ser Ser Asn Tyr
226           85             90             95
227      gtg gtc aaa gat ccc tac aag ttc ctg aac aag gag aag aga gat aag          336
228      Val Val Lys Asp Pro Tyr Lys Phe Leu Asn Lys Glu Lys Arg Asp Lys
229           100            105            110
230      ttc ctg agc agt tac aat tac atc aag gat agc att gac acc gat atc          384
231      Phe Leu Ser Ser Tyr Asn Tyr Ile Lys Asp Ser Ile Asp Thr Asp Ile
232           115            120            125
233      aac ttc gcc aac gat gtc ctg gga tac tac aag atc ctg tcc gag aag          432
234      Asn Phe Ala Asn Asp Val Leu Gly Tyr Tyr Lys Ile Leu Ser Glu Lys
235           130            135            140
236      tac aag agc gat ctg gat agc atc aag aag tac atc aac gat aag cag          480
237      Tyr Lys Ser Asp Leu Asp Ser Ile Lys Lys Tyr Ile Asn Asp Lys Gln
238           145            150            155            160
239      gga gag aac gag aag tac ctg ccc ttc ctg aac aac atc gag acc ctg          528
240      Gly Glu Asn Glu Lys Tyr Leu Pro Phe Leu Asn Asn Ile Glu Thr Leu
241           165            170            175
242      tac aag acc gtc aac gat aag att gat ctg ttc gtg atc cac ctg gag          576
243      Tyr Lys Thr Val Asn Asp Lys Ile Asp Leu Phe Val Ile His Leu Glu
244           180            185            190

```

PAGE: 6

VERIFICATION SUMMARY
PATENT APPLICATION US/09/175,683A

DATE: 01/05/2000
TIME: 15:11:36

Input Set: I175683A.RAW

Line ? Error/Warning

Original Text
